CLAIMS:

1. The nucleotide sequence as shown in Figure 3 or degenerate variants thereof.

2. The nucleotide sequence of claim 1 which has been altered by mutation, deletion or insertion.

3. ORFs derivable from the nucleotide sequence of claim 1 or claim 2; excluding the ORFs identified as y4aL, y4hB, y4hG, y4hH, y4hI, y4nB, y4nC, y4sM, y4vK, y4xA, y4yC, y4yD, y4yE, y4yG in Table 3.

4. ORFs y4aA to y4aS, y4bA to y4bO, y4cA to y4cQ, y4dA to y4dX, y4eA to y4eO, y4fA to y4fR, y4gA to y4gN, y4hA to y4hR, y4iC to y4iR, y4jA to y4jT, y4kA to y4kV, y4lA to y4lS, y4mA to y4mQ, y4nA to y4nM, y4oA to y4oX, y4pA to y4pO, y4qB to y4qK, y4rA to y4rO, y4sA to y4sL, y4sN to y4sO, y4tA to y4tS, y4uA to y4uP, y4vA to y4vS, y4wA to y4wM, y4xA to y4xQ, y4yA to y4yS as identified in Table 3; excluding the ORFs identified as y4aL, y4hB, y4hG, y4hH, y4hI, y4nB, y4nC, y4vK, y4xA, y4yC, y4yD, y4yE, y4yG in Table 3.

5. The ORFs of claim 3 or claim 4 which encode the functions of:

(a) nitrogen fixation,

(b) nodulation,

(c) transportation or permeation,

(d) synthesis and modification of surface poly- or oligosaccharides, lipo-oligosaccharides or secreted oligosaccharide derivatives,

(e) secretion (of proteins or other biomolecules),

(f) transcriptional regulation or DNA-binding,

(g) peptidolysis or proteolysis,

(h) transposition or integration,

(i) plasmid stability, plasmid replication or conjugal plasmid transfer,

(j) stress response (such as heat shock, cold shock or osmotic shock),

(k) chemotaxis,

(l) electron transfer,

(m) synthesis of isoprenoid compounds,

(n) synthesis of cell wall components,

(o) rhizopine metabolism,

(p) synthesis and utilization of amino acids, rhizopines, amino acid derivatives or other biomolecules, or

(q) degradation of xenobiotic compounds,

or which encode:

(r) proteins exhibiting similarities to proteins of amino acid metabolism or related ORFs, or

(r) enzymes (such as oxidoreductase, transferase, hydrolase, lyase, isomerase or ligase).

6. The ORFs of any one of claims 3 to 5 which are under the control of their natural regulatory elements or under the control of analogues to such natural regulatory elements.

7. Intergenic sequences derivable from the nucleotide sequence of claim 1 or claim 2.

8. The intergenic sequences of claim 7 which are regulatory DNA sequences or repeated elements.

9. The intergenic sequences of claim 7 which are ORF-fragments.

10. Mobile elements (insertion elements or mosaic elements) derivable from the nucleotide sequence of claim 1 or claim 2.

11. Proteins expressible from the nucleotide sequences or ORFs of any one of claims 1 to 6.

12. Use of the nucleotide sequences or ORFs of any one of claims 1 to 10 or the proteins of claim 11 in the analysis of the structure, organisation or dynamics of other genomes.

13. Use of the nucleotide sequences or ORFs of any one of claims 1 to 10 or the proteins of claim 11 in:

(a) screening nucleotide sequences,

(b) subcloning nucleotide sequences,

(c) amplifying nucleotide sequences by PCR, or

(d) gene trapping.

14. Use according to claim 13, wherein said nucleotide sequences are coding sequences or non-coding sequences.

15. Use according to claim 14, wherein said coding sequences are regulatory sequences, repeated elements, mosaic sequences or insertion elements.

16. Use according to any one of claims 12 to 15, wherein said nucleotide sequences or ORFs are oligonucleotide primers or hybridization probes.

17. Use of the nucleotide sequences, individual ORFs or groups of ORFs of any one of claims 1 to 10 or the proteins of claim 11 in:

(a) the identification and classification of organisms and their genetic information,

(b) the identification and characterisation of nucleotide sequences, amino acid sequences or proteins,

(c) the transportation of compounds to and from an organism which is host to at least one of said nucleotide sequences, ORFs or proteins,

(d) the degradation and/or metabolism of organic, inorganic, natural or xenobiotic substances in a host organism, or

(e) the modification of the host-range, nitrogen fixation abilities, fitness or competitiveness of organisms.

18. The plasmid comprising the nucleotide sequence of claim 1 or claim 2.

19. A plasmid which harbours at least one ORF of any one of claims 1 to 10 or any degenerate variant thereof or which harbours at least one ORF or any degenerate variant thereof which encodes one or more of the proteins of claim 11.

20. The plasmid of claim 18 or claim 19 produced recombinantly.

21. The plasmid of any one of claims 18 to 20 or any variant thereof produced by mutation, deletion, insertion or inactivation of an ORF, ORFs or groups of ORFs.

22. Use of the plasmid of any one of claims 18 to 21 in:

(a) obtaining a synthetic minimal set of ORFs required for functional *Rhizobium*-legume symbiosis,

(b) the modification of the host-range of rhizobia,

(c) the augmentation of the fitness or competitiveness of *Rhizobium* sp. NGR234 in the soil and its nodulation efficiency on host plants,

(d) the introduction of desired phenotype(s) into host plants using said plasmid as a stable shuttle system for foreign DNA encoding said desired phenotype(s), or

(e) the direct transfer of said plasmid into rhizobia or other microorganisms without using other vectors for mobilization.